

We claim:

1 1. A system for accessing information from a plurality of distributed and
2 heterogeneous data sources each having a plurality of users and applications with a plurality
3 of domain perspectives, the architecture comprising:
4 a user domain module operable to act as an interface with the users and the
5 applications by translating queries from the users and the applications into a form
6 recognizable by an aggregation domain module and by receiving responses from the
7 aggregation domain module and translating the responses into a form recognizable by
8 the users and applications;
9 a generic domain module operable to receive translated queries from the user
10 domain module, translate the queries into a form recognizable to a data source domain
11 module, receive responses from the data source domain module, translate the
12 responses into a form recognizable by the user domain module, and transmit the
13 translated responses to the user domain module; and
14 a data source domain module operable to receive the translated queries from
15 the aggregation domain module, identify target data sources to transmit the queries to,
16 translate the queries specific to the identified data sources, receive responses from the
17 identified data sources, translate the responses, and transmit the translated responses to
18 the aggregation domain module; and
19 a knowledge base comprising a data model of the generic domain, data models of each
20 data source, conceptual terminology translations between the user domains and the generic
21 domain and conceptual terminology translations between the generic domain and the data
22 sources.

1 2. The system according to claim 1, wherein the data model comprises data classes,
2 data classifications, data subclasses, relationships among data classes, and data attributes.

1 3. The system according to claim 1, wherein each conceptual terminology translation
2 comprises a mapping between a term in a first domain and constraints of other terms in the
3 first domain to a term in a second domain and constraints of other terms in the second
4 domain.

1 4. The system according to claim 1, wherein the knowledge base synthesizes user
2 domain conceptual terminology translations, a generic model of the overall domain of
3 discourse, data source domain conceptual terminology translations, and data source models
4 through the user domain module, aggregation module, and the data source module to enable
5 access to a plurality of data sources, and provides knowledge to each of the user domain,
6 aggregation domain, and data source domain.

1 5. The system according to claim 1, wherein the user domain module utilizes concept
2 mapping from the user to generic domains.

1 6. The system according to claim 1, wherein the aggregation domain module
2 comprises a generic data model that describes an overall domain and is described by
3 groupings of data and relationships and classifications among the groupings of data.

1 7. The system according to claim 1, wherein the data source domain module
2 comprises data models of the data sources, and wherein the data models describe groupings of
3 data and relationships and classifications among the groupings of data.

1 8. The system according to claim 1, further comprising:

2 a plurality of user domains each described by concept mappings from user domains to
3 the generic model, wherein the user domain module is operable as an interface between the
4 user applications and the sources by translating queries from the users applications into forms
5 recognizable by the aggregation domain module and operable to translate responses from the
6 aggregation domain module into a form recognizable by the users and the applications.

1 9. The system according to claim 1, wherein the user domain module is operable to

2 receive queries in terminology of the users and applications.

1 10. The system according to claim 1, wherein the user domain module, the

2 aggregation domain module, and the data source domain module are operable to utilize the
3 conceptual terminology translations to translate the queries.

1 11. The system according to claim 1, wherein the aggregation domain uses an overall

2 model of all user domain and database domain models.

1 12. The system according to claim 1, wherein the users and the applications express

2 queries using a high-level request language in the terminology of the user domains and the
3 application domains.

1 13. The system according to claim 1, wherein the user domain module, the

2 aggregation domain module, and the data source domain module utilize the conceptual
3 terminology translations to dissect the queries.

1 14. The system according to claim 1, wherein the data source domain module

2 translates the queries by selecting matching data source contexts.

15. A method for maintaining application and data source autonomy, the method comprising:
providing an intermediate layer that decouples users and applications from underlying data sources.

16. A method for accessing information from distributed heterogeneous sources, the method comprising:
receiving queries from users and applications;
translating the queries into a form from which it can be determined the data sources to address the queries;
determining which data sources the queries should be transmitted to;
formulating data source queries that navigate underlying representations of the data sources in native languages of the data sources;
transmitting the queries to the data sources;
receiving responses to the queries;
aggregating the responses from individual data sources into a single response;
translating the responses into a form recognizable to the users and applications; and
transmitting the responses to the users and applications.

17. The method according to claim 16, wherein the users and applications formulate the queries in a simplified query language.

18. The method according to claim 16, wherein translating the queries comprises:
analyzing the queries by applying concept mappings that describe a mapping of concepts from a user domain to an aggregation domain and from the aggregation domain to a data source domain.

1 19. The method according to claim 16, wherein the queries from the users and the
2 applications are received and translated by a user domain.

1 20. The method according to claim 19, further comprising:
2 receiving the translated queries with a user domain;
3 further translating the queries with an aggregation domain into a form from which it
4 can be determined which data sources to transmit the queries to; and
5 transmitting the further translated queries to a data source domain.

1 21. The method according to claim 20, further comprising:
2 receiving the further translated queries with the data source domain;
3 formulating data source queries that navigate underlying representations of the data
4 sources in native languages of the data sources; and
5 transmitting at least portions of the queries to the data sources.

1 22. The method according to claim 21, wherein the data source domain module
2 receives responses to the queries from the data sources, translates the responses to the queries
3 and transmits the translated responses to the aggregation domain module, wherein the
4 aggregation domain module receives the translated responses, aggregates the data returned
5 from individual data sources, and transmits the translated responses to the user domain
6 module which further translates the responses into a form recognizable to the users and
7 applications.

1 23. The method according to claim 16, wherein determining which data sources to
2 transmit the queries to comprises matching contexts derived from the queries with contexts in
3 the data models.

1 24. A system for accessing information from a plurality of diverse data sources, the
 2 system comprising:
 3 a user domain module, comprising:
 4 a query parser for receiving queries from users and applications, the queries
 5 comprising information to be retrieved, relationships of among the
 6 information, and constraints associated with the information;
 7 an attribute name resolver for translating the queries into a form recognizable
 8 to a generic domain; and
 9 a generic to attribute name converter for translating responses to the queries
 10 into a form recognizable by the applications and users;
 11 an aggregation domain module, comprising:
 12 an attribute name location mapper for receiving translated queries from the
 13 user domain and determining locations of data for responding to the queries;
 14 a location resolver for determining a best location for data when multiple
 15 locations exist; and
 16 a data aggregation function for receiving responses from data sources and
 17 translating data in the responses into a form recognizable by the user domain;
 18 a data source domain module, comprising:
 19 a generic to data source converter for converting query attributes to a form
 20 recognizable to locations where the data is stored;
 21 a query generator to automatically formulate data source queries to navigate
 22 underlying representations of the data sources in native query languages of the
 23 data sources;
 24 a query transmitter for transmitting the formulated queries to the data source
 25 locations; and
 26 a data converter for receiving responses to the formulated queries from the data
 27 storage locations and translating the responses to the terminology of the
 28 generic to attribute name converter; and

29 a knowledge base module, comprising:
 30 a data model of the generic domain, data models of each data source,
 31 conceptual terminology translations between the user domains and the generic
 32 domain and conceptual terminology translations between the generic domain
 33 and the data sources..

1 25. A computer program product for performing an information retrieval process in an
 2 information access system, comprising:
 3 a computer readable medium;
 4 computer program instructions, recorded on the computer readable medium,
 5 executable by a processor, for performing the steps of:
 6 receiving queries from users and applications;
 7 translating the queries into a form from which it can be determined the data sources to
 8 address the queries;
 9 determining which data sources the queries should be transmitted to;
 10 transmitting the queries to the data sources;
 11 retrieving responses to the queries;
 12 aggregating the responses from the data sources into a single response;
 13 translating the responses into a form recognizable to the users and applications; and
 14 transmitting the responses to the users and applications.

1 26. A system for performing an information access process, comprising:
 2 a processor operable to execute computer program instructions; and
 3 a memory operable to store computer program instructions executable by the
 4 processor, for performing the steps of:
 5 receiving queries from users and applications;
 6 translating the queries into a form from which it can be determined the data sources to
 7 address the queries;
 8 determining which data sources the queries should be transmitted to;

- 9 transmitting the queries to the data sources;
- 10 retrieving responses to the queries;
- 11 aggregating the responses from the data sources into a single response;
- 12 translating the responses into a form recognizable to the users and applications; and
- 13 transmitting the responses to the users and applications.

1501-SPL-19E28860